DOCUMENT RESUME

ED 440 979 TM 030 765

AUTHOR Stapleman, Jan
TITLE Standards-Based Accountability Systems. Policy Brief.
INSTITUTION Mid-Continent Regional Educational Lab., Aurora, CO.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 2000-04-00
NOTE 10p.
CONTRACT RJ96006101
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Standards; *Accountability; *Educational Assessment; Elementary Secondary Education; School Districts; State Programs; *Student Evaluation

ABSTRACT

This policy brief summarizes research results and provides guidance regarding decisions associated with school accountability. Unlike previous notions of accountability, a standards-based system examines outputs, such as student performance and graduation rates, as well as inputs like the amount of instructional time or the number of books in the school library. Most states have accomplished the work of adopting statewide content standards and are now engaged in developing assessments that measure what students know and can do in relation to the standards. Student testing is one component of an accountability system; another is rating school performance through methods that incorporate data from assessments and other measures of student success. Once school performance has been measured and rated, issues involving reporting the results surface. After school performance has been measured and rated, it is also essential that struggling schools and districts receive the help they need before they become subject to consequences defined by the accountability system, such as state- or district-imposed sanctions. The granting of rewards or the imposition of sanctions must rest on multiple indicators of school performance. Creating consequences puts teeth into accountability systems, but there is a lack of agreement among experts about their effectiveness and how they should be used. (Contains 15 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 440 979

Policy Brief

McREL April 2000

Standards-Based Accountability Systems

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM030765

MID-CONTINENT RESEARCH FOR EDUCATION AND LEARNING

MID-CONTINENT RESEARCH FOR EDUCATION AND LEARNING

Policy Brief

McREL
April 2000

Standards-Based Accountability Systems

by Jan Stapleman

The notion of accountability isn't new. We all are held accountable in many aspects of daily life. As employees, parents, members of organizations, and citizens, we work to meet certain expectations that others have of us. Our performance and progress are constantly being measured in formal and informal ways.

Those in K-12 education are subject to accountability in similar ways. As states and communities implement measures to hold schools accountable, they confront certain questions: Who should be held accountable? Students? Teachers? Administrators? And for what should each be held accountable?

Traditionally, students have been held accountable for learning through grading systems. Teachers have been held accountable for covering specific content through curricula that are now becoming aligned with established content standards and benchmarks. Principals, superintendents, and other administrators have been held accountable in schools and districts for student test scores and other performance indicators such as graduation rates and student attendance. As accountability systems evolve, states and communities are reevaluating how students, teachers, and administrators are held accountable.

Additional questions remain about how student and school performance and progress should be measured and reported. How should school performance be rated? Who should receive reports on school performance, and for what purpose? What can be done to improve poor performance? How should exemplary performance be recognized? How policymakers answer these questions has a direct bearing on how they shape their state's or district's education accountability system.

Guidance for developing an accountability system

- Standards-based accountability systems work best when all components function together in a coherent fashion to improve student achievement.
- In order to accurately and fairly assess students' progress toward achieving state and local content standards, assessments must be aligned with those content standards.
- Attaching high-stakes consequences to local and statewide testing can motivate schools and students to improve performance, but it also can carry certain risks, including the threat of lawsuits challenging the accuracy and fairness of the tests employed and consequences invoked.
- The best way to evaluate the performance of schools or districts is to consider multiple indicators, such as student achievement, attendance, drop-out rates, and graduation rates.
- Early and ongoing assistance from states and districts can often prevent struggling schools from failing. Resources are often better spent on early intervention, rather than on imposing sanctions after schools have failed.
- Creating consequences such as rewards and sanctions can put teeth into accountability systems, but there is a lack of evidence or agreement among experts about their effectiveness and how they should be used.

BEST COPY AVAILABLE

As accountability measures are put in place, schools, districts, and states play varying roles and have different responsibilities, depending on the way each system has been structured. Each of the 50 states has taken a different approach to holding schools accountable.

States rarely set out to create a new accountability system from whole cloth. A report from Education Commission of the States (1999) noted that components often fall into place in fits and starts, rather than in the logical sequence of developing standards and aligned assessments first. States may implement some components by law and others by regulation. Often components of state systems are not aligned because they were implemented years apart and for different purposes.

In light of the increasing pressures on educators to strengthen, revise, or implement accountability systems, this policy brief attempts to summarize research results and provide guidance regarding decisions associated with school accountability.

Standards-based accountability systems

Unlike past notions of accountability, a standards-based system examines "outputs," such as student performance and graduation rates, as well as "inputs," such as the amount of instructional time and the number of books in the school library.

State and local policymakers and educators all bear responsibility for school success within standards-based accountability systems.

State and local policymakers and educators all bear responsibility for school success within standards-based accountability systems. States hold districts and, in many cases, individual schools accountable for student achievement. In turn, districts and states are responsible for providing ongoing assistance and consequences to struggling schools. Although there is some agreement among education experts on key characteristics of accountability systems, there also is considerable debate about the best way to assure school accountability.

A model of standards-based, state-level accountability systems that has emerged from discussions among experts and an analysis of reform efforts across the nation includes the following components (Education Commission of the States, 1999; Education Week, 1999, p. 9):

- *Aligning standards and assessments*: Congruent state and local content standards and student assessments that are aligned with those standards;
- *Rating school performance*: A rating system that includes multiple indicators such as student achievement, attendance, drop-out rates, and graduation rates;
- *Reporting performance*: A method of reporting school performance to parents, educators, policymakers, and the public, such as school report cards;
- *Providing assistance*: The capacity and will at state and district levels to provide early and ongoing assistance to struggling schools;
- *Creating consequences*: Clearly defined remedies for low-achieving schools and recognition for high-achieving schools.

A focus on accountability also is observed at the federal level, where provisions in the 1994 Elementary and Secondary Education Act (U.S. Department of Education, n.d.) call on states to phase in specific programmatic and reporting requirements by the 2000-2001 school year. A goal of ESEA is to assure that the progress of Title I students in each state be measured with the same assessments used for other students (in at least math and reading), to demonstrate adequate yearly progress by schools. According to that law, if states have a statewide school accountability system, Title I schools must be included in that system.

Aligning standards and assessments

Most states have accomplished the hard work of adopting statewide content standards and have

begun the even more difficult task of developing assessments that accurately measure what students know and are able to do in relation to those standards. In order to implement equitable and accurate assessment, however, states and districts must confront certain questions:

What constitutes fair and appropriate testing?

Standardized tests assess all students in the same, predetermined manner. Critics argue that these tests do not accurately measure in-school student learning because many of the test questions address topics that have not been taught in the classroom. Research studies have shown that some questions on these tests are designed to assess students' intellectual capacity or out-of-school learning, rather than what has been learned in school (Popham, 1999).

Some states and districts use commercially produced, norm-referenced, standardized tests to assess student achievement. Norm-referenced tests measure students' performance against that of other students across the nation. Experts often argue instead for the use of "criterion-referenced" tests, which measure student performance against specific content standards. By the end of 2000, at least 30 states will have developed such tests (Fox, 1999). But criterion-referenced tests have raised different concerns. For example, some parents and policymakers still will want to know how their students compare with students nationwide — information that typically is not provided by criterion-referenced tests (Education Week, 1999, p. 18; Fox, 1999).

Another assessment debate centers on the use of traditional multiple-choice questions versus open-ended questions, portfolios, and performance assessments. Although critics charge that multiple choice questions can't adequately measure complex thinking and problem-solving, nontraditional testing methods have received criticism for being too subjective and not focusing on the basics (Education Week, 1999, p. 16). Further, tests that include constructed response items in addition to multiple choice items are more costly to administer and score.

Budget constraints usually require that state, district, and local policymakers must weigh costs against benefits when selecting assessments. Often it is more cost-effective to purchase a commercial, standardized test. Some experts argue that because such tests are subjected to rigorous validation criteria, standardization procedures, and reliability testing, their results are more useful in comparing, generalizing, and determining levels of attainment of specified standards (Sanders and Horn, 1995). In response to the standards movement, certain commercial test publishers are customizing their tests to fit the content standards and policy objectives of various states, to mixed reviews (Fox, 1999).

A common-sense approach recognizes that no one type of assessment is the best choice in every situation.

Common sense dictates that in order for statewide assessments to measure student learning against state content standards, the tests must be aligned with those standards. Logic also follows that no one type of assessment is the best choice in every situation. Testing within the classroom relies on a variety of methods, including performance assessments and portfolio evaluation (Sanders and Horn, 1995). But many of those methods are difficult and costly to employ when large numbers of students are being tested as part of a statewide accountability system. Using multiple types of assessments is, perhaps, the best way for educators to gain a complete picture of student achievement because they can combine results from commercially available, standardized tests with those from locally developed, alternative assessments.

Who should be tested?

The standards-based reform movement has emphasized that special needs students and English language learners should be included in statewide assessments, based on the belief that schools should be held accountable for the learning of all students. The inclusive nature of

the movement is also supported by legislation. The 1997 IDEA amendments require that all students with disabilities be included in state and district assessments or be given an alternative examination (U.S. Department of Education, 1997). The 1994 ESEA Title I amendments (U.S. Department of Education, n.d.) require that Title I students be tested with the same assessments used for all other students in a state.

The move to include all students in testing creates a dilemma for educators, especially when test results involve high-stakes consequences. For example, testing students in a language they don't understand can produce low, inaccurate test scores. On the other hand, excluding certain groups can produce an inflated overall picture of student performance.

In order to get an accurate measure of learning among all their student groups, states, districts, and schools must test all students except those with the most severe disabilities, providing appropriate accommodations for students with disabilities or students who are learning English. Examples of reasonable accommodations may include such provisions as administering the test in a separate location, or in more than one session, or in the student's native language or Braille (Landau, Vohs, and Romano, 1998). Test results can be interpreted with more accuracy by reporting the scores of student subgroups in addition to overall student performance (Linn, 1998).

What are the risks of high-stakes testing?

Experts continue to debate the wisdom of employing high-stakes tests — tests that carry significant consequences for schools, educators, or students. For schools, those consequences may involve the amount of future funding or the threat of sanctions. For educators, they may include reassignment or termination. For students, they may affect the ability to graduate or advance to the next grade.

Many educators and parents credit their districts' use of high-stakes testing for prompting students to get serious about learning. A survey conducted by Public Agenda in conjunction with *Quality Counts* '99, found that 68 percent of high school students queried said exit exams "make them work harder" (Education Week, 1999, pp. 53-54).

But sometimes high-stakes tests produce undesirable and unintended consequences, such as teaching the test or excluding some students from testing (Fuhrman, 1999). Tying assessments to students' graduation or promotion can prompt students to drop out or increase the number of years necessary to graduate (Education Week, 1999, pp. 55-56). High-stakes testing also can invite court challenges to the accuracy and fairness of the measurement tools (Barnes, 1999; Institute for the Study of Educational Policy, 1994; Phillips, 1993). When schools, districts, and states evaluate their accountability systems, it is a good idea to examine not only the expected, positive effects, but also any unintended, negative consequences (Linn, 1998).

Rating school performance

Another essential component of accountability systems is a method of rating school performance that incorporates data generated from assessments and other measures of student success. The data should relate directly to learning and school improvement goals. In order to implement a rating system, schools, districts, and states must determine what outcomes to evaluate, define satisfactory performance, and decide whether or not to give credit for improved performance. Some states and districts also have a system of ranking schools in relation to one another.

What outcomes should be evaluated?

Rating schools according to student performance on a single test is an inherently unreliable way to measure school success. State and district policymakers can minimize criticism of test adequacy and fairness by examining a broader set of success indicators, rather than relying only on student achievement measures.

In addition to test scores, some state accountability systems incorporate measures such as graduation rates, drop-out rates, and attendance. Other factors that have been linked by research with improvements in test scores —

and over which schools have some control — include climate, course-taking patterns, levels of parent involvement, and the proportion of teachers who are teaching subjects in which they majored in college or have been certified to teach (Education Week, 1999, p. 33).

What is satisfactory performance?

Defining satisfactory performance is largely a subjective judgment. State and district policymakers can avoid the appearance of arbitrariness by defining satisfactory school performance in terms that are clear and understandable to students, parents, and the public (Fuhrman, 1999). It is important to promote public understanding by explaining the standards-setting process and providing examples of items and adequate performance at each level.

Credit for improved performance?

Should poorly funded schools or schools that serve large numbers of students who arrive ill-prepared to learn be held accountable to the same performance standards as well-funded schools or schools with students who are better prepared? In order to hold schools accountable for factors within their control, some accountability systems focus on measuring progress in student achievement. But critics cite many examples of schools serving disadvantaged populations where achievement is high. They point out that continually accepting modest growth from low-performing schools might mean some students never get the education they need to compete as adults.

One solution is to hold schools accountable for both the level of student achievement and progress in student achievement (Fuhrman, 1999). Setting both long-term and short-term goals for all schools allows for differences in student preparedness during early assessments but ultimately requires greater growth rates from low-performing schools (Linn, 1998).

Reporting performance

Once school performance has been measured and rated, how should the results be reported to parents, educators, policymakers, and the public? Many states publish school report cards.

The purpose of school report cards is to make the results of school improvement efforts public by reporting student achievement and progress made. The value of school report cards depends on whether they include information that is meaningful to parents, policymakers, and the public in general. The value of information currently included in school report cards varies widely among states. Many fail to include crucial factors, such as those noted above under "What outcomes should be evaluated?" (Education Week, 1999, p. 33).

Some critics charge that the reports are a waste of time and money if they end up gathering dust on bookshelves (Education Week, 1999, p. 36). Most experts agree that report cards are most useful when they include pertinent information about student and school progress, guide future improvement efforts, and are widely disseminated to parents and the public through mailings, the media, and postings on state department of education Web sites.

Assistance to low-performing schools

Once school performance has been measured and rated, it is essential that struggling schools or districts receive needed assistance before they are subject to consequences such as state- or district-imposed sanctions. Some accountability systems seem to presume that schools have the capacity to improve student performance if they simply can be motivated to do so. But merely imposing sanctions prescribed by an accountability system is unlikely to unleash hidden capacity in failing schools. Experience shows that many failing schools educate disproportionate numbers of disadvantaged students and are in need of support — from technical assistance to professional development to hands-on help from expert educators or state representatives (Education Week, 1999, p. 38).

Many experts caution that the achievement gap between privileged and underprivileged students will persist until all children have access to the qualified teachers and adequate resources they need in order to excel (Linn, 1998). In short, all students must be given the opportunity to learn.

States and districts are finding it is to their benefit to provide ongoing technical, professional, and/or financial assistance to struggling schools early, before sanctions are necessary (Education Week, 1999, p. 38; Regional Educational Laboratory Network, 1998).

Creating consequences

Once states have measured student performance against standards and rated schools on multiple measures of success, they must confront the fact that some schools will emerge as highly successful and others may not measure up, even with assistance. Some state accountability systems set forth consequences, including monetary or nonmonetary rewards for highly successful schools and/or sanctions for failing schools. Policymakers in these states, and in other states in the process of implementing such measures, feel consequences put teeth into accountability systems, which otherwise would have little impact. Nevertheless, there is little evidence that punitive consequences lead to improved outcomes.

Rewards

Many states point to positive results from programs rewarding high-performing schools and their teachers. Supporters believe that bestowing public recognition (and, perhaps, cash) upon successful schools and teachers is an effective incentive (Education Week, 1999, pp. 61-64). ECS (1999) identified the absence of rewards as one of three design deficiencies revealed in its survey of accountability systems in the 50 states.

Performance pay for teachers should be tied to explicit standards for teachers while performance awards for schools should be tied to schoolwide improvements in student achievement.

—Allan Odden

But critics say that merit pay programs, where teachers are rewarded for performance rather than for seniority, discourage collegiality by pitting educators against each other. Allan Odden (1999) argued that performance pay for teachers should be tied to explicit standards for teachers while performance awards for schools should be tied to schoolwide improvements in student achievement. But offering financial bonuses to schools with high student scores may actually discourage highly qualified teachers from working in the most challenging schools and may encourage "teaching the test" (Education Week, 1999, pp. 62-63).

If rewards are used, state and district policymakers should base them on indicators that are valid and reliable and disseminate them in a way that is perceived as fair. Once a monetary rewards program is in place, the program's funding must be sustained over time if the accountability system is to be taken seriously by educators (ECS, 1999).

Sanctions

Some state policymakers consider reporting school performance as an end in itself, believing the embarrassment of being publicly designated as low performing will often motivate school personnel to rally their troops and find ways to improve performance (Education Week, 1999, p. 38). Policymakers in other states believe that failing schools need consequences, such as loss of accreditation, loss of state funding, state takeovers, closing, or reconstitution (which often involves replacing school principals, teachers, and staff).

State and district leaders recognize that providing early and ongoing assistance to struggling schools can sometimes prevent having to impose sanctions. But, even with assistance, some schools may not have the leadership, teacher expertise, or other resources needed to overcome the momentum of a downward spiral in student achievement. In such cases, more extreme measures may be necessary to turn student performance around.

Sanctions can produce unintended consequences, however, especially since they tend to fall disproportionately on schools attended by poor and minority students. Of the schools listed by states as failing, more than half are in urban

areas, four in 10 have minority enrollments greater than 90 percent, and three in four are designated as high-poverty schools. Such schools usually lack the resources of better-funded schools and employ younger, less experienced teachers (Education Week, 1999, p. 38).

In the final analysis, many states are reluctant to follow through on imposing severe sanctions like academic takeover or reconstitution. In a takeover situation, the state often finds itself grappling with most of the same problems (and, perhaps, the same lack of capacity) that departing administrators faced. On the other hand, once a state threatens to impose sanctions, it is important to follow through; failure to do so damages the credibility of its accountability system.

Conclusion

Designing standards-based school accountability systems is a complex process. Although the various components often are implemented over time and in response to varying events and conditions, local and state-level accountability systems work best when all components function together in a coherent fashion to improve student achievement.

It also is a challenging task to design or select effective assessments. The best way for educators to obtain a clear picture of student achievement is through the use of multiple types of tests. In order to accurately and fairly assess students' progress toward achieving state and local content standards, the assessments must be aligned with those content standards.

Considerable debate exists about the wisdom of attaching high-stakes consequences to local and statewide testing results. The practice can motivate schools and students to improve performance, but it also can carry certain risks, including the threat of lawsuits challenging the accuracy and fairness of the tests employed.

The most accurate and fair way to evaluate the performance of schools or districts is to consider multiple indicators, such as student achievement, attendance, drop-out rates, and graduation rates. Rating systems that rely on the results of a single test are far more likely to be unfair and inaccurate.

The value of reports on school success is determined by the relevance of the information on student and school progress they provide and how broadly they are disseminated to parents and other stakeholders.

Early and ongoing assistance from states and districts can often prevent struggling schools from failing. However, even with assistance, some schools may not succeed and may require certain sanctions. Other schools will excel, raising the question of whether their efforts should be rewarded. Creating consequences such as rewards and sanctions can put teeth into accountability systems, but there is a lack of evidence or agreement among experts about their effectiveness and how they should be used.

ADDITIONAL RESOURCES

- For more information on interpreting and meeting the requirements of the 1994 ESEA Title I amendments, visit ED's Web site at www.ed.gov/offices/OESE/StandardsAssessment/overview.html
- A McREL policy brief on high-stakes testing will be published soon. Additional guidance on the complex legal questions surrounding high-stakes testing can be found in a guide developed by the North Central Regional Educational Laboratory (Phillips, 1993).

Jan Stapleman is communications coordinator at McREL.

Sources

Barnes, E. (1999). Schools should use caution in high-stakes testing. Education Daily, 32(77), 1.

Education Commission of the States. (1999). Education accountability systems in 50 states. Denver, CO: Author.

Education Week. (1999). Quality counts '99: Rewarding results, punishing failure. Education Week, 18(17).

Fox, J. (1999). Old-style tests may hamper school reform, experts say. Education Daily, 32(70), 1.

BEST COPY AVAILABLE

Fuhrman, S. (1999). The new accountability. CPRE Policy Briefs, RB-27-January 1999. Philadelphia: Consortium for Policy Research in Education.

Institute for the Study of Educational Policy. (1994). Likely issues in linking certificates of mastery and high school diplomas: Legal lessons from minimum competency testing. CSL Policy Briefs. University of Washington. [On-line]. Available: http://csl.wednet.edu/Web%20page/9%20Resources/policy%20briefs/policy%20brief_7.html

Landau, J. K., Vohs, J. R., & Romano, C. A. (1998). All kids count. Boston: Federation for Children with Special Needs.

Linn, R. L. (1998). Standards-based accountability: Ten suggestions. CRESST Policy Brief. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Odden, A. (1999, June 27). Set standards for teacher performance. Denver Post [On-line]. Available: http://www.denverpost.com/opinion/odden0627.htm

Phillips, S. E. (1993). Legal implications of high-stakes assessment: What states should know. Chicago: North Central Regional Educational Laboratory.

Popham, W. J. (1999). Why standardized tests don't measure educational quality. Educational Leadership, 56(6), 8.

Regional Educational Laboratory Network. (1998). Curriculum reform: What state officials say works. Denver, CO: Mid-continent Regional Educational Laboratory.

Sanders, W. L., & Horn, S. P. (1995). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. Education Policy Analysis Archives, 3(6). [On-line]. Available: http://olam.ed.asu.edu/epaa/v3n6.html

U. S. Department of Education. (n.d.) Guidance on standards, assessments, and accountability. [On-line]. Available: http://www.ed.gov/offices/OESE/StandardsAssessment/assess.html

U. S. Department of Education, Office of Special Education Programs. (1997). The individuals with disabilities act amendments of 1997: Curriculum, Module 6. Washington, DC: Author.

This publication is based on work sponsored wholly, or in part, by the Office of Educational Research and Improvement (OERI), U.S. Department of Education, under contract number RJ96006101. The content of this publication does not necessarily reflect the views of OERI, the department, or any other agency of the U.S. government.

McREL

Tim Waters, executive director & president
Jana Caldwell, director of communications
Jan Stapleman, editor
Jeanne Deak, graphic designer

Mid-continent Research for Education and Learning
2550 South Parker Road, Suite 500
Aurora, Colorado 80014
phone 303.337.0990
fax 303.337.3005
Visit us at our Web site: www.mcrel.org

ACCOUNTABILITY INFORMATION FROM McREL FOR EDUCATION POLICYMAKERS!

Mid-continent Research for Education and Learning
2550 S. Parker Road, Suite 500
Aurora, CO 80014